



SEQUENCE LISTING

RECEIVED
OCT 26 2000TECH CENTER 1600/2900
RECEIVED

OCT 30 2000

TECH CENTER 1600/2900

(1) GENERAL INFORMATION:

- (i) APPLICANT: von Schaewen, Antje Dr. rer. nat.
- (ii) TITLE OF INVENTION: Plant gntI sequences and the use thereof for the production of plants having reduced or lacking N-acetylglucosaminyl transferase I (GnTI) activity
- (iii) NUMBER OF SEQUENCES: 6
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: von Schaewen, Antje Dr. rer. nat.
 - (B) STREET: Natruperstrasse 169a
 - (C) CITY: Osnabrueck
 - (D) COUNTRY: Germany
 - (E) ZIP: D-49076
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA
- (A) APPLICATION NUMBER: 09/591,466
 - (B) FILING DATE: 06/09/2000
 - (C) CLASSIFICATION: 536

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1669 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: Solanum tuberosum
 - (B) STRAIN: Desiree
 - (D) DEVELOPMENTAL STAGE: Sink organ
 - (F) TISSUE TYPE: Mesophyll
 - (G) CELL TYPE: Leaf cells
- (vii) IMMEDIATE SOURCE:
- (A) LIBRARY: Lambda ZAP II (Eco RI)
 - (B) CLONE: gntI-Al(K)
- (ix) FEATURE:
- (A) NAME/KEY: misc_feature
 - (B) LOCATION: 659..667
 - (D) OTHER INFORMATION: /function= "Asn codon in this context is a potential glycosylation site"
/product= "N-glycosylation consensus sequence"

```

1      /phenotype= "N-glycans modulate protein
2      properties"
3      /standard_name= "N-glycosylation site"
4      /label= pot-CHO
5      /note= "GnTI-coding sequences from animals do not
6      contain this feature"
7
8      (ix) FEATURE:
9          (A) NAME/KEY: CDS
10         (B) LOCATION:53..1393
11         (C) IDENTIFICATION METHOD: experimental
12         (D) OTHER INFORMATION:/codon_start= 53
13             /function= "initiates complex N-glycans on
14             secretory glycoproteins"
15             /EC_number= 2.4.1.101
16             /product=
17             "beta-1,2-N-acetylglucosaminyltransferase I"
18             /evidence= EXPERIMENTAL
19             /gene= "cgl"
20             /standard_name= "gntI"
21             /label= ORF
22             /note= "first gntI sequence from potato (unpublished)"
23
24         (ix) FEATURE:
25             (A) NAME/KEY: 5'UTR
26             (B) LOCATION:15..52
27
28         (ix) FEATURE:
29             (A) NAME/KEY: 3'UTR
30             (B) LOCATION:1394..1655
31
32         (ix) FEATURE:
33             (A) NAME/KEY: CDS
34             (B) LOCATION:80..139
35             (D) OTHER INFORMATION:/function= "membrane anchor (amino
36             acids 10-29)"
37             /product= "hydrophobic amino acid stretch in GnTI"
38             /standard_name= "membrane anchor of a type II
39             Golgi protein"
40             /note= "identified by comparison with GnTI sequences
41             from animals"
42
43         (ix) FEATURE:
44             (A) NAME/KEY: misc_feature
45             (B) LOCATION:1..14
46             (D) OTHER INFORMATION:/function= "used for cloning the
47             cDNA library in Lambda ZAPII"
48             /product= "EcoRI/NotI-cDNA adapter"
49             /number= 1
50
51         (ix) FEATURE:
52             (A) NAME/KEY: misc_feature
53             (B) LOCATION:1656..1669
54             (D) OTHER INFORMATION:/product= "EcoRI/NotI-cDNA adapter"
55             /number= 2
56
57         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
58
59      GAATTCGCGG CCGCCTGAGA AACCTCGAA TTCAATTCG CATTGGCAG AG ATG      55
60
61                                         Met
62                                         1
63

```

1	AGA GGG AAC AAG TTT TGC TTT GAT TTA CGG TAC CTT CTC GTC GTG GCT	103
2	Arg Gly Asn Lys Phe Cys Phe Asp Leu Arg Tyr Leu Leu Val Val Ala	
3	5 10 15	
4		
5	GCT CTC GCC TTC ATC TAC ATA CAG ATG CGG CTT TTC GCG ACA CAG TCA	151
6	Ala Leu Ala Phe Ile Tyr Ile Gln Met Arg Leu Phe Ala Thr Gln Ser	
7	20 25 30	
8		
9	GAA TAT GTA GAC CGC CTT GCT GCT GCA ATT GAA GCA GAA AAT CAT TGT	199
10	Glu Tyr Val Asp Arg Leu Ala Ala Ile Glu Ala Glu Asn His Cys	
11	35 40 45	
12		
13	ACA AGT CAG ACC AGA TTG CTT ATT GAC AAG ATT AGC CAG CAG CAA GGA	247
14	Thr Ser Gln Thr Arg Leu Leu Ile Asp Lys Ile Ser Gln Gln Gln Gly	
15	50 55 60 65	
16		
17	AGA GTA GTA GCT CTT GAA GAA CAA ATG AAG CAT CAG GAC CAG GAG TGC	295
18	Arg Val Val Ala Leu Glu Glu Gln Met Lys His Gln Asp Gln Glu Cys	
19	70 75 80	
20		
21	CGG CAA TTA AGG GCT CTT GTT CAG GAT CTT GAA AGT AAG GGC ATA AAA	343
22	Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile Lys	
23	85 90 95	
24		
25	AAG TTA ATC GGA GAT GTG CAG ATG CCA GTG GCA GCT GTA GTT GTT ATG	391
26	Lys Leu Ile Gly Asp Val Gln Met Pro Val Ala Ala Val Val Val Met	
27	100 105 110	
28		
29	GCT TGC AGT CGT ACT GAC TAC CTG GAG AGG ACT ATT AAA TCC ATC TTA	439
30	Ala Cys Ser Arg Thr Asp Tyr Leu Glu Arg Thr Ile Lys Ser Ile Leu	
31	115 120 125	
32		
33	AAA TAC CAA ACA TCT GTT GCA TCA AAA TAT CCT CTT TTC ATA TCC CAG	487
34	Lys Tyr Gln Thr Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser Gln	
35	130 135 140 145	
36		
37	GAT GGA TCA AAT CCT GAT GTA AGA AAG CTT GCT TTG AGC TAT GGT CAG	535
38	Asp Gly Ser Asn Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Gly Gln	
39	150 155 160	
40		
41	CTG ACG TAT ATG CAG CAC TTG GAT TAT GAA CCT GTG CAT ACT GAA AGA	583
42	Leu Thr Tyr Met Gln His Leu Asp Tyr Glu Pro Val His Thr Glu Arg	
43	165 170 175	
44		
45	CCA GGG GAA CTG GTT GCA TAC TAC AAG ATT GCA CGT CAT TAC AAG TGG	631
46	Pro Gly Glu Leu Val Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys Trp	
47	180 185 190	
48		
49	GCA TTG GAT CAG CTG TTT CAC AAG CAT AAT TTT AGC CGT GTT ATC ATA	679
50	Ala Leu Asp Gln Leu Phe His Lys His Asn Phe Ser Arg Val Ile Ile	
51	195 200 205	
52		
53	CTA GAA GAT GAT ATG GAA ATT GCT GCT GAT TTT TTT GAC TAT TTT GAG	727
54	Leu Glu Asp Asp Met Glu Ile Ala Ala Asp Phe Phe Asp Tyr Phe Glu	
55	210 215 220 225	
56		
57	GCT GGA GCT ACT CTT CTT GAC AGA GAC AAG TCG ATT ATG GCT ATT TCT	775
58	Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile Ser	
59	230 235 240	
60		
61	TCT TGG AAT GAC AAT GGA CAA AGG CAG TTC GTC CAA GAT CCT GAT GCT	823
62	Ser Trp Asn Asp Asn Gly Gln Arg Gln Phe Val Gln Asp Pro Asp Ala	
63	245 250 255	

```

1
2 CTT TAC CGC TCA GAC TTT TTT CCT GGT CTT GGA TGG ATG CTT TCA AAA 871
3 Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser Lys
4 260 265 270
5
6 TCA ACT TGG TCC GAA CTA TCT CCA AAG TGG CCA AAG GCT TAC TGG GAT 919
7 Ser Thr Trp Ser Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp Asp
8 275 280 285
9
10 GAC TGG CTA AGG CTG AAA GAA AAT CAC AGA GGT CGA CAA TTT ATT CGC 967
11 Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile Arg
12 290 295 300 305
13
14 CCA GAA GTT TGC AGA ACG TAC AAT TTT GGT GAG CAT GGT TCT AGT TTG 1015
15 Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser Leu
16 310 315 320
17
18 GGG CAG TTT TTT AAG CAG TAT CTT GAG CCA ATT AAG CTA AAT GAT GTC 1063
19 Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp Val
20 325 330 335
21
22 CAG GTT GAT TGG AAG TCA ATG GAC CTA AGT TAC CTT TTG GAG GAC AAC 1111
23 Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp Asn
24 340 345 350
25
26 TAT GTG AAA CAC TTT GGC GAC TTG GTT AAA AAG GCT AAG CCC ATC CAC 1159
27 Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile His
28 355 360 365
29
30 GGA GCT GAT GCT GTT TTG AAA GCA TTT AAC ATA GAT GGT GAT GTG CGT 1207
31 Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val Arg
32 370 375 380 385
33
34 ATT CAG TAC AGA GAC CAA CTA GAC TTT GAA GAT ATC GCT CGA CAG TTT 1255
35 Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asp Ile Ala Arg Gln Phe
36 390 395 400
37
38 GGC ATT TTT GAA GAA TGG AAG GAT GGT GTA CCA CGG GCA GCA TAT AAA 1303
39 Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr Lys
40 405 410 415
41
42 GGG ATA GTA GTT TTC CGG TTT CAA ACA TCT AGA CGT GTG TTC CTT GTT 1351
43 Gly Ile Val Val Phe Arg Phe Gln Thr Ser Arg Arg Val Phe Leu Val
44 420 425 430
45
46 TCC CCT GAT TCT CTT CGA CAA CTT GGA GTT GAA GAT ACT TAG 1393
47 Ser Pro Asp Ser Leu Arg Gln Leu Gly Val Glu Asp Thr *
48 435 440 445
49
50 CGAAGATATG ATTGGAGCCT GAGCAACAAT TTAGACTTAT TTGGTAGGAT ACATTTGAAA 1453
51
52 GAGCTGACAC GAAAAGTATG ACTACCAGTA GCTACATGCA ACATTTTAAT GTTAATGGAA 1513
53
54 GGAACCCACT GCTTATTGTT GGAATGGATG AATCATCACC ACATCCTATT ATTCAAGTTT 1573
55
56 ACAAACATAA AGAGGAAAATG TTGCCCTATA AAAACAAATT TTTTGTTTCT AAGAAGGAAC 1633
57
58 GTTACGATTA TGAGCAACTT TGGCGGCCGC GAATTC 1669
59
60 (2) INFORMATION FOR SEQ ID NO: 2:
61
62 (i) SEQUENCE CHARACTERISTICS:
63 (A) LENGTH: 446 amino acids

```

```

1          (B) TYPE: amino acid
2          (D) TOPOLOGY: linear
3
4          (ii) MOLECULE TYPE: protein
5          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
6
7  Met Arg Gly Asn Lys Phe Cys Phe Asp Leu Arg Tyr Leu Leu Val Val
8      1          5          10          15
9
10 Ala Ala Leu Ala Phe Ile Tyr Ile Gln Met Arg Leu Phe Ala Thr Gln
11      20          25          30
12
13 Ser Glu Tyr Val Asp Arg Leu Ala Ala Ile Glu Ala Glu Asn His
14      35          40          45
15
16 Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Lys Ile Ser Gln Gln Gln
17      50          55          60
18
19 Gly Arg Val Val Ala Leu Glu Glu Gln Met Lys His Gln Asp Gln Glu
20      65          70          75          80
21
22 Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile
23      85          90          95
24
25 Lys Lys Leu Ile Gly Asp Val Gln Met Pro Val Ala Ala Val Val Val
26      100         105         110
27
28 Met Ala Cys Ser Arg Thr Asp Tyr Leu Glu Arg Thr Ile Lys Ser Ile
29      115         120         125
30
31 Leu Lys Tyr Gln Thr Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser
32      130         135         140
33
34 Gln Asp Gly Ser Asn Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Gly
35      145         150         155         160
36
37 Gln Leu Thr Tyr Met Gln His Leu Asp Tyr Glu Pro Val His Thr Glu
38      165         170         175
39
40 Arg Pro Gly Glu Leu Val Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys
41      180         185         190
42
43 Trp Ala Leu Asp Gln Leu Phe His Lys His Asn Phe Ser Arg Val Ile
44      195         200         205
45
46 Ile Leu Glu Asp Asp Met Glu Ile Ala Ala Asp Phe Phe Asp Tyr Phe
47      210         215         220
48
49 Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile
50      225         230         235         240
51
52 Ser Ser Trp Asn Asp Asn Gly Gln Arg Gln Phe Val Gln Asp Pro Asp
53      245         250         255
54
55 Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser
56      260         265         270
57
58 Lys Ser Thr Trp Ser Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp
59      275         280         285
60
61 Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile
62      290         295         300
63

```

```

1  Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser
2  305                      310                      315                      320
3
4  Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp
5                      325                      330                      335
6
7  Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp
8                      340                      345                      350
9
10 Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile
11                      355                      360                      365
12
13 His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val
14                      370                      375                      380
15
16 Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asp Ile Ala Arg Gln
17 385                      390                      395                      400
18
19 Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr
20                      405                      410                      415
21
22 Lys Gly Ile Val Val Phe Arg Phe Gln Thr Ser Arg Arg Val Phe Leu
23                      420                      425                      430
24
25 Val Ser Pro Asp Ser Leu Arg Gln Leu Gly Val Glu Asp Thr
26                      435                      440                      445
27
28 (2) INFORMATION FOR SEQ ID NO: 3:
29
30 (i) SEQUENCE CHARACTERISTICS:
31 (A) LENGTH: 1737 base pairs
32 (B) TYPE: nucleic acid
33 (C) STRANDEDNESS: double
34 (D) TOPOLOGY: linear
35
36 (ii) MOLECULE TYPE: cDNA to mRNA
37
38 (iii) HYPOTHETICAL: NO
39
40 (iv) ANTI-SENSE: NO
41
42 (vi) ORIGINAL SOURCE:
43 (A) ORGANISM: Nicotiana tabacum
44 (B) STRAIN: Samsun NN
45 (D) DEVELOPMENTAL STAGE: Sink organ
46 (F) TISSUE TYPE: Mesophyll
47 (G) CELL TYPE: Leaf cells
48
49 (vii) IMMEDIATE SOURCE:
50 (A) LIBRARY: Lambda ZAP II (Eco RI)
51 (B) CLONE: gntI-A9(T)
52
53 (ix) FEATURE:
54 (A) NAME/KEY: misc feature
55 (B) LOCATION: 733..741
56 (D) OTHER INFORMATION: /function= "Asn codon in this
57 context is a potential glycosylation site"
58 /product= "N-glycosylation consensus sequence"
59 /phenotype= "N-glycans modulate protein
60 properties"
61 /standard_name= "N-glycosylation site"
62 /label= pot-CHO
63 /note= "GnTI sequences from animals do not contain this

```

```

1         feature"
2
3     (ix) FEATURE:
4         (A) NAME/KEY: CDS
5         (B) LOCATION:127..1467
6         (C) IDENTIFICATION METHOD: experimental
7         (D) OTHER INFORMATION:/codon_start= 127
8             /function= "initiates complex N-glycans on
9                 secretory glycoproteins"
10            /EC_number= 2.4.1.101
11            /product=
12                "beta-1,2-N-acetylglucosaminyltransferase I"
13            /evidence= EXPERIMENTAL
14            /gene= "cgl"
15            /standard_name= "gntI"
16            /label= ORF
17            /note= "first gntI sequence from tobacco (unpublished)"
18
19     (ix) FEATURE:
20         (A) NAME/KEY: 5'UTR
21         (B) LOCATION:15..126
22
23     (ix) FEATURE:
24         (A) NAME/KEY: 3'UTR
25         (B) LOCATION:1468..1723
26
27     (ix) FEATURE:
28         (A) NAME/KEY: CDS
29         (B) LOCATION:154..213
30         (D) OTHER INFORMATION:/function= "membrane anchor (amino
31             acids 10-29)"
32             /product= "hydrophobic amino acid stretch in GntI"
33             /standard_name= "membrane anchor of a type II
34                 golgi protein"
35
36     (ix) FEATURE:
37         (A) NAME/KEY: misc_feature
38         (B) LOCATION:1..14
39         (D) OTHER INFORMATION:/function= "use for cloning the
40             cDNA library in Lambda ZAPII"
41             /product= "EcoRI/NotI-cDNA adapter"
42             /number= 1
43
44     (ix) FEATURE:
45         (A) NAME/KEY: misc_feature
46         (B) LOCATION:1724..1737
47         (D) OTHER INFORMATION:/product= "EcoRI/NotI-cDNA adapter"
48             /number= 2
49
50
51     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
52
53     GAATTCGCGG CCGCCATTGA CTTGATCCTA ACTGAACAGG CAAAGTAAAT CCAGCGATGA      60
54
55     AACACTCATA ACTGAACACT GAGAGACTAT TCGCTTTCTC CTAAAGCCTT CAATCGAATT      120
56
57     CGCACG ATG AGA GGG AAC AAG TTT TGC TGT GAT TTC CGG TAC CTC CTC      168
58         Met Arg Gly Asn Lys Phe Cys Cys Asp Phe Arg Tyr Leu Leu
59             450             455             460
60
61     ATC TTG GCT GCT GTC GCC TTC ATC TAC ACA CAG ATG CGG CTT TTT GCG      216
62     Ile Leu Ala Ala Val Ala Phe Ile Tyr Thr Gln Met Arg Leu Phe Ala
63             465             470             475

```

4

	720	725	730	
1				
2				
3	TAC TGG GAC GAC TGG CTA AGA CTC AAA GAG AAT CAC AGA GGT CGA CAA			1032
4	Tyr Trp Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln			
5	735	740	745	
6				
7	TTT ATT CGC CCA GAA GTT TGC AGA ACA TAT AAT TTT GGT GAG CAT GGT			1080
8	Phe Ile Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly			
9	750	755	760	765
10				
11	TCT AGT TTG GGG CAG TTT TTC AAG CAG TAT CTT GAG CCA ATT AAA CTA			1128
12	Ser Ser Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu			
13		770	775	780
14				
15	AAT GAT GTC CAG GTT GAT TGG AAG TCA ATG GAC CTT AGT TAC CTT TTG			1176
16	Asn Asp Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu			
17		785	790	795
18				
19	GAG GAC AAT TAC GTG AAA CAC TTT GGT GAC TTG GTT AAA AAG GCT AAG			1224
20	Glu Asp Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys			
21		800	805	810
22				
23	CCC ATC CAT GGA GCT GAT GCT GTC TTG AAA GCA TTT AAC ATA GAT GGT			1272
24	Pro Ile His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly			
25		815	820	825
26				
27	GAT GTG CGT ATT CAG TAC AGA GAT CAA CTA GAC TTT GAA AAT ATC GCA			1320
28	Asp Val Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asn Ile Ala			
29		830	835	840
30				
31	CGG CAA TTT GGC ATT TTT GAA GAA TGG AAG GAT GGT GTA CCA CGT GCA			1368
32	Arg Gln Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala			
33		850	855	860
34				
35	GCA TAT AAA GGA ATA GTA GTT TTC CGG TAC CAA ACG TCC AGA CGT GTA			1416
36	Ala Tyr Lys Gly Ile Val Val Phe Arg Tyr Gln Thr Ser Arg Arg Val			
37		865	870	875
38				
39	TTC CTT GTT GGC CAT GAT TCG CTT CAA CAA CTC GGA ATT GAA GAT ACT			1464
40	Phe Leu Val Gly His Asp Ser Leu Gln Gln Leu Gly Ile Glu Asp Thr			
41		880	885	890
42				
43	TAA CAAAGATATG ATTGCAGGAG CCCGGGCAAA ATTTTGTACT TATTGGGTAG			1517
44	*			
45				
46				
47	GATGCATCGA GCTGACACTA AACCATGATT TTACCAGTTA CATAACAACGT TTTAATGTTA			1577
48				
49	TACGGAGGAG CTCACTGTTC TAGTGTGAA GGGATATCGG CTTCTTAGTA TTGGATGAAT			1637
50				
51	CATCAACACA ACCTATTATT TTAAGTGTTT AGAACATAAA GAGGAAATGT AGCCCTGTAA			1697
52				
53	AGACTATACA TGGGACCATC ATAATCGCGG CCGCGAATTC			1737
54				
55				
56	(2) INFORMATION FOR SEQ ID NO: 4:			
57				
58	(i) SEQUENCE CHARACTERISTICS:			
59	(A) LENGTH: 446 amino acids			
60	(B) TYPE: amino acid			
61	(D) TOPOLOGY: linear			
62				
63	(ii) MOLECULE TYPE: protein			

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

1      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
2
3      Met Arg Gly Asn Lys Phe Cys Cys Asp Phe Arg Tyr Leu Leu Ile Leu
4      1          5          10
5
6      Ala Ala Val Ala Phe Ile Tyr Thr Gln Met Arg Leu Phe Ala Thr Gln
7      20          25          30
8
9      Ser Glu Tyr Ala Asp Arg Leu Ala Ala Ala Ile Glu Ala Glu Asn His
10     35          40          45
11
12     Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Gln Ile Ser Leu Gln Gln
13     50          55          60
14
15     Gly Arg Ile Val Ala Leu Glu Glu Gln Met Lys Arg Gln Asp Gln Glu
16     65          70          75          80
17
18     Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile
19     85          90          95
20
21     Lys Lys Leu Ile Gly Asn Val Gln Met Pro Val Ala Ala Val Val Val
22     100         105         110
23
24     Met Ala Cys Asn Arg Ala Asp Tyr Leu Glu Lys Thr Ile Lys Ser Ile
25     115         120         125
26
27     Leu Lys Tyr Gln Ile Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser
28     130         135         140
29
30     Gln Asp Gly Ser His Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Asp
31     145         150         155         160
32
33     Gln Leu Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val His Thr Glu
34     165         170         175
35
36     Arg Pro Gly Glu Leu Ile Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys
37     180         185         190
38
39     Trp Ala Leu Asp Gln Leu Phe Tyr Lys His Asn Phe Ser Arg Val Ile
40     195         200         205
41
42     Ile Leu Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp Phe Phe
43     210         215         220
44
45     Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile
46     225         230         235         240
47
48     Ser Ser Trp Asn Asp Asn Gly Gln Met Gln Phe Val Gln Asp Pro Tyr
49     245         250         255
50
51     Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser
52     260         265         270
53
54     Lys Ser Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp
55     275         280         285
56
57     Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile
58     290         295         300
59
60     Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser
61     305         310         315         320
62
63     Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp

```

```

1           325           330           335
2
3 Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp
4           340           345           350
5
6 Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile
7           355           360           365
8
9 His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val
10          370           375           380
11
12 Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asn Ile Ala Arg Gln
13          385           390           395           400
14
15 Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr
16           405           410           415
17
18 Lys Gly Ile Val Val Phe Arg Tyr Gln Thr Ser Arg Arg Val Phe Leu
19           420           425           430
20
21 Val Gly His Asp Ser Leu Gln Gln Leu Gly Ile Glu Asp Thr
22          435           440           445
23
24

```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1854 base pairs
- (B) TYPE: Nucleotide
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: Columbia
- (D) DEVELOPMENTAL STAGE: Mature plants
- (F) TISSUE TYPE: All tissues

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Lambda Uni-ZAP (EcoRI/XhoI) and
Lambda ACT (XhoI)
- (B) CLONE: pBSK-Ara-GntI-full #8

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1185..1193
- (D) OTHER INFORMATION: /function= "Asn Codon is a
potential glycosylation site"
/product= "Consensus sequence for
N-glycosylation"
/phenotype= "N glycans modulate
protein characteristics"
/standard_name= "N glycosylation site"
/label= pot-CHO
/note= "absent in animal GnTI sequences"

(ix) FEATURE:

```

1      (A) NAME/KEY: CDS
2      (B) LOCATION:135..1469
3      (C) IDENTIFICATION METHOD: experimental
4      (D) OTHER INFORMATION:/codon_start= 135
5          /function= "initiates complex N glycans on
6          secretory glycoproteins"
7          /EC_number= 2.4.1.101
8          /product=
9          "beta-1,2-N-acetyl glucosaminyl transferase I"
10         /evidence= EXPERIMENTAL
11         /gene= "cgl"
12         /standard_name= "gntI"
13         /label= ORF
14         /note= "first gntI sequence from Arabidopsis
15         (unpublished)"
16
17     (ix) FEATURE:
18         (A) NAME/KEY: 5'UTR
19         (B) LOCATION:19..134
20
21     (ix) FEATURE:
22         (A) NAME/KEY: 3'UTR
23         (B) LOCATION:1470..1848
24
25     (ix) FEATURE:
26         (A) NAME/KEY: CDS
27         (B) LOCATION:157..215
28         (D) OTHER INFORMATION:/function= "membrane anchor
29         (amino acids 8-27)"
30         /product= "hydrophobic amino-acid region in
31         GnTI"
32         /standard_name= "membrane anchor of a Type II
33         Golgi protein"
34         /note= "identified by comparison with animal GnTI
35         sequences "
36
37     (ix) FEATURE:
38         (A) NAME/KEY: misc_feature
39         (B) LOCATION:1..18
40         (D) OTHER INFORMATION:/function= "for preparation
41         of a cDNA library in Lambda ACT"
42         /product= "XhoI-cDNA-Adaptor"
43         /number= 1
44
45     (ix) FEATURE:
46         (A) NAME/KEY: misc_feature
47         (B) LOCATION:1849..1854
48         (D) OTHER INFORMATION:/product= "XhoI-cDNA-Adaptor"
49         /number= 2
50
51
52     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
53
54     CTCGAGGCCA CGAAGGCCAC CGTTTTTGTT ATAACGAACG ACACCGTTTC AAACAAC TTC      60
55
56     CTTATTAGCT AGCTCCCTCC CGGCGGCAAA CACCAGAAGA TCCACCGCTT TTGATCTGGT      120
57
58     TGTTTGTCGT CGAT ATG GCG AGG ATC TCG TGT GAC TTG AGA TTT CTT CTC      170
59         Met Ala Arg Ile Ser Cys Asp Leu Arg Phe Leu Leu
60         1          5          10
61
62     ATC CCG GCA GCT TTC ATG TTC ATC TAC ATC CAG ATG AGG CTT TTC CAG      218
63     Ile Pro Ala Ala Phe Met Phe Ile Tyr Ile Gln Met Arg Leu Phe Gln

```

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 0
- L1
- L2
- L3
- L4
- L5
- L6
- L7
- L8
- L9
- M0
- M1
- M2
- M3
- M4
- M5
- M6
- M7
- M8
- M9
- N0
- N1
- N2
- N3
- N4
- N5
- N6
- N7
- N8
- N9
- O0
- O1
- O2
- O3
- O4
- O5
- O6
- O7
- O8
- O9

1 (A) LENGTH: 444 amino acids
 2 (B) TYPE: Amino acid
 3 (D) TOPOLOGY: Linear
 4
 5 (ii) MOLECULE TYPE: Protein
 6 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
 7
 8 Met Ala Arg Ile Ser Cys Asp Leu Arg Phe Leu Leu Ile Pro Ala Ala
 9 1 5 10 15
 10
 11 Phe Met Phe Ile Tyr Ile Gln Met Arg Leu Phe Gln Thr Gln Ser Gln
 12 20 25 30
 13
 14 Tyr Ala Asp Arg Leu Ser Ser Ala Ile Glu Ser Glu Asn His Cys Thr
 15 35 40 45
 16
 17 Ser Gln Met Arg Gly Leu Ile Asp Glu Val Ser Ile Lys Gln Ser Arg
 18 50 55 60
 19
 20 Ile Val Ala Leu Glu Asp Met Lys Asn Arg Gln Asp Glu Glu Leu Val
 21 65 70 75 80
 22
 23 Gln Leu Lys Asp Leu Ile Gln Thr Phe Glu Lys Lys Gly Ile Ala Lys
 24 85 90 95
 25
 26 Leu Thr Gln Gly Gly Gln Met Pro Val Ala Ala Val Val Val Met Ala
 27 100 105 110
 28
 29 Cys Ser Arg Ala Asp Tyr Leu Glu Arg Thr Val Lys Ser Val Leu Thr
 30 115 120 125
 31
 32 Tyr Gln Thr Pro Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser Gln Asp
 33 130 135 140
 34
 35 Gly Ser Asp Gln Ala Val Lys Ser Lys Ser Leu Ser Tyr Asn Gln Leu
 36 145 150 155 160
 37
 38 Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val Val Thr Glu Arg Pro
 39 165 170 175
 40
 41 Gly Glu Leu Thr Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys Trp Ala
 42 180 185 190
 43
 44 Leu Asp Gln Leu Phe Tyr Lys His Lys Phe Ser Arg Val Ile Ile Leu
 45 195 200 205
 46
 47 Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp Tyr Phe Glu Ala
 48 210 215 220
 49
 50 Ala Ala Ser Leu Met Asp Arg Asp Lys Thr Ile Met Ala Ala Ser Ser
 51 225 230 235 240
 52
 53 Trp Asn Asp Asn Gly Gln Lys Gln Phe Val His Asp Pro Tyr Ala Leu
 54 245 250 255
 55
 56 Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Lys Arg Ser
 57 260 265 270
 58
 59 Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp Asp Asp
 60 275 280 285
 61
 62 Trp Leu Arg Leu Lys Glu Asn His Lys Gly Arg Gln Phe Ile Ala Pro
 63 290 295 300

1
 2 Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser Leu Gly
 3 305 310 315 320
 4
 5 Gln Phe Phe Ser Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp Val Thr
 6 325 330 335
 7
 8 Val Asp Trp Lys Ala Lys Asp Leu Gly Tyr Leu Thr Glu Gly Asn Tyr
 9 340 345 350
 10
 11 Thr Lys Tyr Phe Ser Gly Leu Val Arg Gln Ala Arg Pro Ile Gln Gly
 12 355 360 365
 13
 14 Ser Asp Leu Val Leu Lys Ala Gln Asn Ile Lys Asp Asp Asp Arg Ile
 15 370 375 380
 16
 17 Arg Tyr Lys Asp Gln Val Glu Phe Glu Arg Ile Ala Gly Glu Phe Gly
 18 385 390 395 400
 19
 20 Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Thr Ala Tyr Lys Gly
 21 405 410 415
 22
 23 Val Val Val Phe Arg Ile Gln Thr Thr Arg Arg Val Phe Leu Val Gly
 24 420 425 430
 25
 26 Pro Asp Ser Val Met Gln Leu Gly Ile Arg Asn Ser
 27 435 440
 28